11 The Analysis of Competition by Field Experiments

Tony Underwood

11.1 Introduction

Controversy continues to rage about the relative importance of interspecific competition among species as a major determinant of the structure of natural communities. [See Chapter 6 and the arguments developed in Strong *et al* (1984b), and the issue of the *American Naturalist* (**122**, 1983) devoted to the topic.] Much of this debate is still a heated discussion of the recognition of patterns in nature that might be explained either as a result of competition in the past, or because of present day competitive processes. It is worth pointing out that such debate does not directly shed light on the problem of whether competition is actually important or not — it only generates heat about whether we can observe and document patterns of distribution and abundance of species that are consistent with competitive processes. For example, Simberloff (1978) and coworkers (Connor & Simberloff 1979; Strong *et al* 1979) have argued cogently that *post hoc* interpretations based on former competition among species are awry, and should be rethought, when the patterns of distribution of species on islands are not distinguishable from random colonization and occupancy of different islands. To rebut this line of reason with statements about its incorrectness (e.g. Diamond & Gilpin 1982; Gilpin & Diamond 1982) is to make no concrete progress in demonstrating the underlying importance of competition. Competitive interactions may or may not explain nonrandomness in species' distributions on islands. Indeed, intense competition may even have led to apparent randomness of distributions of species (although hypotheses to explain this are currently sparse). In both cases, the existence or nonexistence of nonrandom patterns does not, and cannot, demonstrate the importance or unimportance of competition as the process leading to the modern pattern.

The demonstration of patterns in space and time that are consistent with competitive interactions is obviously stronger evidence for the importance of these processes than would be the absence of such patterns. The patterns themselves are, however, of no value except to raise hypotheses about competition — they cannot (even where there is no argument about their existence) simultaneously be used as tests of the importance of competition. Dayton (1973) and Schroder and Rosenzweig (1975) have provided very convincing demonstrations of the need for well-planned field experiments to prevent totally spurious conclusions about processes in natural communities.

Correctly, therefore, there has been considerable investment of time, energy and ingenuity into the planning and carrying out of field experiments designed to eliminate alternative hypotheses, and thereby to support competition as an important process to explain present patterns of distribution and abundance of organisms. Several recent reviews have drawn attention to the great number of experiments on competition between organisms in the field (Birch 1979; Lawton & Hassell 1981; Lawton & Strong 1981; Connell 1983; Schoener 1983b). Experimental studies of competitive interactions in natural populations are scarcely new (Jackson 1981), and the reviews listed indicate a great abundance and diversity of studies on a wide range of organisms. What is not clear from these reviews is how good is the evidence. What sort of experiments are claimed to be successful in producing unequivocal evidence for competitive interactions? Are the available studies of such good quality that there is no longer any room for doubt about the widespread importance of competitive interactions? Connell (1983) has suggested that field experiments are often regarded as the *ne plus ultra* of ecological research, and has pointed out that poor experiments are often worse than useless because the results are often accepted without question. Schoener (1983b) opted for the position that experimental studies are complementary to observational studies (rather than the position that observations are useful for the proposal of explanatory models that can then be tested by properly designed experiments), but nevertheless concluded that field experiments had revealed much about interspecific competition.

I am less concerned about the outcome of experiments on interspecific competition than about their validity. Although the reviews already mentioned tend to the view that much is known, that many satisfactory experiments are pointing to some widespread advances and generalizations about the nature of interspecific competition, a more cautious question to ask is how good are field experiments on competition? This is not an idle question. There have been, in parallel with attempts to synthesize the current wad of experimentally acquired knowledge about competition, several reviews of the nature of biological experimentation, and the pitfalls and problems of much of the work. Dayton (1979), Dayton and Oliver (1980) and Underwood and Denley (1984) have pointed out serious deficiencies in the development of hypotheses and the lack of consideration of alternative viewpoints in the design, implementation and interpretation of many field experiments. Widespread problems with complex sets of experimentally acquired biological data have been revealed in the analyses discussed by Underwood (1981) and Hurlbert (1984). Is the experimental study of interspecific competition somehow better than other areas of ecological experimentation? How well will the currently available studies stand up when analysed by less personally involved historians of ecology at some point in the future when today's hypotheses and theory might be as interesting and relevant as Ptolemy's?

In the search for new theory today's is found wanting, will future ecologists view today's experiments excitedly, as an abandoned gold-mine to be reopened, or as the ideological equivalent of a used-car lot, littered with the rusty remains of vehicles of self-advancement, now abandoned by careless speedsters down the vanished highways of former ecological theory?

The present analysis is therefore of the experiments as experiments, not as contributors to theory. How well conceived have the experiments been? How might designs be improved? Are the available experimental studies as useful as recent reviews have suggested?

11.1.1 Sources of studies

I have benefited from the spade-work of Connell (1983) and Schoener (1983b), and have used their extensive reviews of experimental studies of competition as the source of material for the present purposes. Previous reviews have already selected from among the enormous number of papers available, using various criteria. Connell, in particular, emphasized the variance in quality of much of the available work, and excluded some papers that Schoener considered. All the studies listed by these two authors have not been covered, but, instead, the sample was limited to 57 papers published in *Ecology* (1960–1982), 16 in *Oecologia* (1975–1982 only), 10 in the *Journal of Animal Ecology* (1960–1982), 11 in the *Journal of Ecology* (1970–1982), and one paper in the *American Naturalist* (1981) (it was the only paper in that journal cited by Schoener as containing descriptions of experiments) — a total of 95 publications. The sample is by no means random, but it covers all the papers in Schoener from the journals and years indicated. This arbitrary number was examined to determine the hypotheses being tested, the nature and design of the experiments, the type of controls used, the degree of replication in space and/or time, the type of analysis of the data and the interpretations made by the authors. My selection biases the types of study to those currently in vogue with those arbiters of fashion who referee and edit the chosen journals, but may not be unrepresentative (and was considered to be largely representative by the other reviewers mentioned). Some 63 of the papers examined for the present work were also included in Connell's (1983) discussion; nine of the papers were also included in Birch's (1979) less extensive review. The decision to limit the sample to 95 papers (rather than any round, but still arbitrary, number) was determined by two things. First, the sample examined had sufficient evidence of some of the problems with field experimentation that a larger sample was not needed to gain some impression of the difficulties being experienced by the authors and, presumably, other readers. Secondly, as will be devel-

oped below, despair and tedium set in, in equal proportions, before I reached the larger individual papers in *Ecological Monographs* that were to make up 100 publications.

Several serious problems were encountered in the papers examined, including some studies that were partially incomprehensible, some that were not about competition, some that were not really experiments, and some that were not particularly in the field. Others were lacking in sensible controls and/or replication. Still others were designed to be confounded so that interesting comparisons, and crucial tests of important hypotheses were not possible, were invalid, or only indirectly interpretable. These topics are illustrated in the body of this chapter. Fortunately, however, some studies were of great value as indications of the successful application of hypothetico-deductive logic and simple principles of experimental design. These are also discussed below. Finally, some specific designs of experiments are described to illustrate procedures that may help eliminate some of the problems encountered here.

11.2 Non-experiments sometimes not about interspecific competition and not always in the field

11.2.1 Studies not about competition

In one of the studies examined, the experimental procedures did not really examine competitive interactions (regardless of the claims of the original author or the subsequent reviewers). Benke (1978) discovered that survivorship of early season species of dragonfly was unaffected by the arrival of later, smaller species. The late species were, however, deleteriously affected by the presence of the earlier species, because the early species ate the later arrivals. Benke (1978) considered this predation to be an extreme form of interference competition. I, like Connell (1983), wish to preserve the fundamental difference between predation and competition. The concept of competitive interactions is only important when the interactions occur as a result of a shortage of some critical resource required by the competing animals (Birch 1957). Predation by one species on another is not a competitive process, even if the prey might, if left untouched by the predator, compete with the predator for other food resources. In Benke's (1978) experiments, competition for food (small animals) had been hypothesized. The food resources were not in short supply in the enclosures and, as pointed out by Benke, dragonflies rarely die of shortage of food. Competitive interactions should not be invoked where resources are apparently adequate for all the consumers present.

11.2.2 Studies not in the field

Studies of organisms in the field are also problematic because it is often difficult to retain experimental densities or combinations of species in natural areas without the use of artificial enclosures (fences, cages, etc.). These installations may bring artefacts into the results that prevent sensible conclusions being drawn about the demonstrated competition and whether it occurs under natural conditions. Apart from noting that experimental enclosures might introduce diseases as artificial sources of mortality (leading to increased recognition of apparent competition where mortality is increased when densities are large) or may prevent the natural attrition of experimental organisms because predators are excluded from the experimental plots (thus making a competitive interaction appear more important than it might ever be in nature), no value judgements are offered about the accuracy of the match between experiments and nature. To do the experiments at all requires some interference with the natural train of events (see later for more discussion of proper controls in field experiments).

There are, however, investigations where laboratory experiments are appropriate for the questions being addressed, such as Wrobel *et al*'s (1980) experiments on salamanders. Such studies clearly should not be included in discussions of field experiments [although this study was included in Schoener's (1983b) list of field experiments].

Several studies have been made under conditions where the densities of the organisms investigated were raised well above naturally occurring densities. It is not clear from these papers that the ranges of densities investigated actually would occur in nature. Thus, Istock (1973) in a well-designed analysis of interactions among various species of corixids in a pond, increased the densities above those occurring naturally ("to overload the system somewhat at the beginning"), although the starting densities and weights used for stocking the experimental ponds were not obvious in the paper.

The analysis of competition between blue tits (*Parus caerules*) and great tits (*P. major*) is often done using nest-boxes instead of natural roosting or nest-sites (Dhondt & Eyckerman 1980; Minot 1981). In Minot's (1981) study, this presumably raised the densities of breeding pairs above those normally encountered when artificial nest-boxes are not available. Similarly, Werner and Hall (1977, 1979) in their study of interactions between freshwater fish, used densities somewhat greater than those occurring naturally. In experiments on intertidal grazing gastropods (Underwood 1978) I replaced dead animals at approximately fortnightly intervals through the experiments. This is entirely unnatural — adult snails and limpets would not appear rapidly in the system under natural, field conditions. In defence of these experiments, however, the

mortality of the superior competitor *Nerita atramentosa* was very slight in cages with the limpet *Cellana tramoserica*; consequently, replacements of *Nerita* were very few, and could not have altered the conclusions from the experiments.

Finally, some of the classic experiments on barnacles by Connell (1961), which have deservedly been widely quoted as an excellent example of a successful field experiment on competition, were not done under natural conditions. To determine the nature of interactions between a high-shore species of British barnacle (*Chthamalus stellatus*) with the barnacle *Balanus balanoides*, that settled in great numbers lower on the shore, Connell mapped naturally settled *Chthamalus* and also moved stones with *Chthamalus* to various levels on the shore and removed all of the *Balanus* from one half of each stone, leaving the other half intact as a control. The stones were transplanted from higher levels, where *Chthamalus* had settled, to lower levels, because the barnacles did not settle lower than mean tide-level. Thus, the intense interference from *Balanus* reduced the densities of *Chthamalus* on all stones within and outside the observed natural range of distribution of the upper-shore species. *Balanus* did not eliminate the *Chthamalus* from stones at low levels (see also Underwood & Denley 1984). Connell's (1961b) great contribution was to demonstrate that competitive interactions could be investigated by manipulative experiments in the field. It does not, however, demonstrate the importance of competition as a determinant of the natural distribution of the upper species of barnacle, because competition with *Balanus* did not eliminate *Chthamalus* even in the unnatural situation where they were moved below their normal level of distribution on the shore.

All these papers clearly reveal much about competitive interactions between species. It is nevertheless arguable whether these particular studies should be used to determine the incidence and importance of competition in nature. They may not actually be about competition occurring in nature at natural densities of the participants.

11.2.3 'Natural experiments' interpreted as manipulative studies

General agreement appears to exist amongst the proponents of ecological field experiments that so-called 'natural' experiments are of little or no value in eliminating conflicting hypotheses about processes such as competition. [See the excellent discussion in Connell (1974).] A 'natural' experiment is one where the hypothesis concerns the effects of a putative competing species (species B) on the biology or demography of a target species (A). Such an investigation commonly consists of comparisons of the mortality, growth, migration, reproduction, etc., of species A in areas

where the supposed competitor (B) is naturally absent, with the biology of species A in supposedly similar areas where the other species (B) is present. This type of comparison cannot eliminate the alternative hypothesis that the biology of species A differs between the two areas for reasons other than the presence of species B (and thus not as a result of any competitive interaction at all). The very nature of the 'experiment' makes it likely that noncompetitive hypotheses based on differences from one habitat to another are more important. The two types of area are, after all, always chosen *because* they are different (i.e. one set of study areas does not contain one species). For this reason, it would be more sensible if ecologists ceased to dignify such comparisons with the label 'experiment' and thus ceased to accord them similar ranking with experiments designed specifically to test hypotheses about competition. The latter are capable of distinguishing among alternative hypotheses that are confounded in the natural comparison.

In contrast, alternative hypotheses based on potential differences among the two sorts of habitat are easily eliminated in the common form of experimental manipulation [what Hurlbert (1984) called the "unnatural" experiment] where the potential competing species (B) is removed from some areas it shares with A, and the performance of A is then monitored in comparison with control, untouched situations.

This is well-known, and scarcely worth reiterating, except that the literature examined contained several natural experiments, that were accepted without question by some or all of the previous reviewers. For example, Gross and Werner (1982), in an excellent study of the colonization of established fields by various biennial plants, had numerous replicated plots sown with seed of each of four species. Pre-emptive competition from the established plants was suggested to be an important agent preventing such seeds from becoming established in some study areas. This was inferred partially from the fact that seeds germinated and survived better in a field where more bare ground was available than in two other fields. The fields were, however, chosen for study because they had not been cultivated for differing periods (1, 5 and 15 years). Competition was deduced from the comparative success of seeds in the three fields, and differences in germination and subsequent survivorship could just as easily have been due to *any* factors correlated with the differences in age (soil condition, nutrition, wind scour, etc.), not just the differences in cover of established plants. The potential competitors (the existing plants) were not manipulated in this part of the study. In a study of experimental introductions of teasel *Dipsacus sylvestris* into old-fields, Werner (1977) manipulated the colonization by the teasel, but did no experimental alterations of the existing plants, nor the litter on the ground, yet concluded that differences from one experimental field to another in the success of the invasion by teasels were due to pre-emptive

competition by the existing plants, or because the resource of bare space was not available where extensive litter had formed under the existing coverage of plants. Again, the results of an otherwise perfectly good manipulative experiment that revealed patterns of difference from one field to another were explained by processes of competition inferred from a natural experimental comparison.

McLay (1974) provided a very clear example of a natural experiment for the supposed competitive interactions between duck-weed *Lemna perpusilla* and the weed *Potamogeton pectinatus* in a pond. Duckweed did not grow well over *Potamogeton*; according to McLay this was because of alterations of the pH of the water due to photosynthetic activity of the *Potamogeton* (and thus a form of chemical inhibition or competition). The major test of this hypothesized competition was during one summer when a dense front of duckweed uncharacteristically grew out over a bed of *Potamogeton*. The pH under the duckweed then altered, which was attributed to the shading of the lower weeds by the *Lemna*, decreasing their photosynthetic rate and thereby altering the pH to more favourable conditions for the duckweed. Again, there was no manipulation of abundance of either weed. The uncharacteristic growth of *Lemna* during that summer could well have been due to environmental conditions that also resulted in altered pH of the water, or simultaneously resulted in changes in the photosynthetic rate of *Potamogeton*, leading to altered pH, regardless of the presence of the duckweed.

Classic natural experiments were found in other papers examined. McClure (1980) contrasted the performance of two species of scale insects on hemlock by examining the insects on six plants that naturally had one species only, six with the other and six that naturally had both.

Some studies involved natural, nonexperimental comparisons in part of the work. For example, Davis (1973) suggested that golden-crowned sparrows *Zonotrichia atricapilla* were unaffected by competition from juncos *Junco hyemalis*, because the distribution of the sparrows did not alter during a period when juncos were naturally sparse. Such comparisons are equally suspect whether they claim to demonstrate or fail to demonstrate supposed competition. Similarly, Waser (1978) studied competition by plants for the resource of humming-birds (to act as pollinators). The flowering success of *Delphinium nelsoni* flowering before the presumed competitor (*Ipomopsis aggregata*) was compared with that of *Delphinium* flowering simultaneously with the *Ipomopsis*. If pollinators were sparse, the two species might be in competition for the attentions of the humming-birds, and thus flower less successfully than if the other species of plant were not also flowering. The flowering success of these two sorts of *Delphinium* may have differed for all sorts of reasons, including different times of the year, or any form of genetic difference in response to selection for flowering early or late in the season, etc. Again,

this is a 'natural' experiment. Part of Morris and Grant's (1972) analysis of competition between rodents consisted of experimental removals of *Clethrionomys*, to determine the effects on *Microtus*. There was no replication, as noted by the authors (and see below), but they considered the experiment to have been repeated by examination of a population of *Microtus* after the natural die-off of *Clethrionomys* during one winter. This is obviously a natural experiment; winter conditions causing the deaths of *Clethrionomys* might have been advantageous to the subsequent well-being of *Microtus*. In all cases, the study of competition was considerably better served by subsequent, manipulative experiments that produced far less equivocal results.

One study, however, demonstrated a more subtle form of natural experiment, because they were wholly embedded in an otherwise valid experimental manipulation. Peckarsky and Dodson (1980) examined competition for prey among stoneflies (*Acroneuria* and *Megarcys*) in streams. Responses to prey were examined by recording migrations of the predators into and out of cages where prey were included or absent. Competition among the predators was apparently examined by comparing the colonization by other predatory species into cages that happened to contain a stonefly at the end of the previous experiments, with cages that had no predator inside. This is a 'natural' experiment because the presence or absence of stoneflies was not a result of any experimental manipulation. This is in no way different from a more obviously 'natural' experiment examining colonization into areas naturally containing some putative competitor compared with similar areas where that species was naturally absent. Rates of colonization into these cages may have differed in many ways, regardless of the previous occupancy by a putative competing predator. If other predatory species had different habitat preferences from those shown by the stoneflies, they would be *expected* to colonize different cages from those containing a stonefly. A better experiment would have been to remove stoneflies from half of the cages and then compare colonization into those with (controls) and those without (experimental) stoneflies. In this way, there would have been, on average, no potential intrinsic differences between cages containing a stonefly, and those without.

Thus, in the 95 papers examined, one was not really about competition, a further six different papers were arguably not about competition occurring in any natural manner under field conditions, and no less than eight included unmanipulated, correlative studies (so-called 'natural' experiments). Together, these studies constitute some 16% of the 95 papers examined. Uncritical acceptance of these papers as examples of successful field experiments about competition seriously overestimates the availability of such evidence.

11.3 Replication and controls in experiments

11.3.1 Lack of replication

There should be no further debate about the need for replication of experimental units, particularly in field studies where many variables are known to be uncontrolled, and many are obviously uncontrollable. It is inconceivable that only the desired, experimentally manipulated variable of interest will actually differ between any two pieces of the world (Connell 1974), particularly where these two pieces are arbitrarily chosen as control and experimental sites from among the possible sites available. Hurlbert (1984) has made an important contribution to this problem of the design of field experiments by pointing out the pitfalls of inadequate replication. In particular, he has drawn attention to the perils of 'pseudoreplication', where the incorrect level or type of experimental unit is replicated. Consider an experiment to investigate the effects on a target species (A) of removal of a hypothesized competing species (B) from one area (called the experimental area), by making comparisons with the ecology, survival, biology, behaviour, and so forth, of species A in the experimental area with that in an untouched area (called a control). It does not matter how many replicated readings, samples, traps, quadrats, etc., are actually monitored in each of these areas, the difference between the two plots is still confounded. Differences might be (as often hoped or simply asserted by experimenters) due to the experimental treatment (i.e. the removal of species B), or due to intrinsic differences between the plots at the start of, or during, the experiment (as is often denied, without any evidence, by the experimenters). The only sensible method for distinguishing between the two possible explanations for differences between the two experimental areas (control and experimental) would be to replicate the units, by having several experimental plots and several controls. In this manner, the potential intrinsic differences between plots can be considered as negligible, because they will have been averaged over the randomly chosen control plots, and averaged equally over the randomly chosen experimental plots. For further consideration of this, consult any sensible book about statistical procedures and random sampling [and such ecological advice as that in Eberhardt (1976) and Green (1979)]. Attempts to overcome the problems associated with lack of replication are sometimes made by testing for differences in various variables between the two study sites before experimental manipulation. This achieves nothing, even in cases where some considerable armoury of sophisticated statistical procedures is used, and many potentially important variables are examined. Lack of a significant difference in the rainfall,

plant cover, numbers of each species, rate of attack by predators, etc., is not the same thing as *no difference* in these variables between the two areas. As Hurlbert (1984) has pointed out, such a lack of significant difference in any applied test is simply a result of inadequate sampling of the variables in the two areas. It is inconceivable that two randomly chosen areas will be identical in all the variables of interest or importance. Such prior tests are of *no* value in determining that two study sites will remain the same, except for the manipulated variable of interest, during the course of an experiment, particularly when the experiment may run for several weeks or years. The only experiments that can have any claim to being unequivocal are those that are replicated in space or time, by the inclusion of several independent control and experimental units into the design of the experiment. This is, of course, well known and should need no further discussion for present-day ecologists ["No one would now dream of testing the response to a treatment by comparing two plots, one treated and the other untreated" (Fisher & Wishart 1930)].

It comes as a surprise therefore to discover how many experiments on competition [15 papers (16%) of those examined] had no replication. Some experiments were simply done with no replication (Pontin 1960; Werner & Hall 1977, 1979; Peterson 1979; Bertness 1981a; Grace & Wetzel 1981). In some studies, there were replicates for some of the experimental treatments, but not for others. For example, Duggins (1980, 1981), in two totally separate series of experiments, had no replicates of the control plots (from which putative competitors were not removed). Menge (1972), in contrast, had data from several control plots, but no replicated areas of experimentally increased and decreased densities of starfish. In still other studies, there was pseudoreplication in various of the comparisons being made. For example, Hixon (1980) compared the effects of reef fish (*Embiotica lateralis* and *E. jacksoni*) by removing each species from one reef, and leaving a third, untouched reef as a control; this experiment was repeated at two depths. Several replicate counts of the numbers of fish of each species on each reef were then made. These are, however, pseudoreplicates in that the differences perceived between reefs are not unequivocally attributable to the experimental removals of the fish. Similarly, Montgomery (1981) removed one species of rodent (*Apodemus*) from one experimental grid, removed another species from a second grid, and left a third grid untouched as a control, despite the wide range of density of one of the species (*A. flavicollis*) throughout the study area (p. 130 of his paper). In this study, numerous replicated traps in rows, with replicated trap-points, traps, and nights of sampling were irrelevant to the validity of attributing differences between the three grids to the removals of species from two of them. Paine (1980) examined the effects of chitons (*Katharina*) on other grazing molluscs by removing chitons from a single experimental area,

and comparing results with a single control plot. Replicated quadrats within each site were used to draw conclusions about the nature of interactions between these grazers. All that such pseudoreplicates can do is to indicate the existence of patterns of difference between experimental units (sites, grids, etc.). No amount of such replication can help to attribute the differences specifically to the experimental treatment.

In some studies [a further 11 papers (12%) of those sampled], there were replicates of the experimental treatments, but these were not considered as such in any comparisons among experimental treatments, or between experimental plots and controls. Thus, Haven (1973) repeated his experimental removals of each of two species of limpets in three separate experiments. The repeated experiments were not considered as replicates in any analysis. In this type of repeated, but unreplicated, experiment, differences among the repeats are usually ignored. Sometimes, differences among repeated experimental treatments must be attributed to various predetermined differences among the repeated sites or times of sampling. Lubchenco (1980) removed the thallus, or thallus and crust of the alga *Chondrus crispus* to determine its competitive effect on the distribution of other algae, *Fucus* spp. In the experiment, there was one plot cleared of the thallus, one cleared of thallus and crust, and a single control (untouched) plot. The experiment was repeated in two different sites (chosen originally because they were subject to different exposure to wave-action, and therefore the repeated experimental treatments were not actually replicates).

The excuses offered for lack of proper replication are as numerous and diverse as the types of experiments themselves. Some authors are, however, willing to concede that the degree of replication of their experiments was inadequate. Thus, Inger and Greenberg (1966) discussed the serious shortcomings of their experiments, and the possible effects of these on the results obtained. They were, however, working in a climate of considerable political uncertainty that allowed no subsequent experimentation. Montgomery (1981) lacked time to do better experiments. Morris and Grant (1972) considered the limitations of their experiments on rodents. Undoubtedly, lack of money or manpower must have limited the number of experimental plots that could be managed in some studies. Nevertheless, none of these causes of lack of replication can make up the serious deficiencies in unreplicated experimental work. For the purposes of the present analysis of experiments, those without replicates are useless. They do not form a proper test of any of the stated hypotheses about competition, because they are not designed in order to falsify many other hypotheses that could explain the patterns in the experimental data. In keeping with the aims of modern ecological experimentation, unreplicated studies do not really constitute valid experiments.

Some types of organisms are obviously much more difficult to study

than others, because of mobility, size, rareness, size of territory, home-range or sparsity of natural resources. It is therefore more likely that investigation of these organisms (e.g. birds and mammals, in contrast to sessile marine invertebrates) will end up unreplicated. Examples of studies on small mammals and birds are those of Davis (1973), Chappell (1978), Minot (1981). It is difficult, however, to avoid the conclusion that lack of attention to the design of experiments and the need for replication are more important factors in the appalling state of much current experimentation that is the intrinsic difficulty of the subject matter. For example, replicated studies on small mammals have been done by Redfield *et al* (1977), Inouye (1981) and Joule and Jameson (1972). None of these authors reported any particular difficulty in achieving their experimental requirements. In contrast, many studies of rodents are unreplicated (Montgomery 1981; Abramsky & Sellah 1982) or are repeated with different treatments, in different places or at different times without any adequate replication in the repeats (e.g. DeLong 1966; Grant 1971; Price 1978; Holbrook 1979). Similarly, experiments on salamanders can easily be replicated as in the elegant studies by Hairston (1980b, 1981) and Wilbur (1972); there were no replicates in Jaeger's (1971) experiments, although they were repeated in different habitats. Competition between lizards was studied in replicated areas by Smith (1981); Tinkle's (1982) experiments on members of the same genera had no replication. There is usually no particular difficulty in manipulating replicate areas of a seashore as shown by replicated studies on limpets (Stimson 1970; Underwood 1978; Creese & Underwood 1982) and sessile invertebrates (Taylor & Littler 1982). Some studies of reef fish are replicated (e.g. Williams 1981) as are some studies of plants in oldfields (Allen & Forman 1976).

An overwhelming conclusion about field experiments on competitive interactions is that the authors apparently choose whether or not to include replicates in the experiments more or less at whim. This occurs despite the widely acknowledged need for replication in field experiments, and the lack of convincing justifications of experimental results when absence of replication leads to potential ambiguities of interpretations. If the simplest requirement for the design of an experiment had to be met by field ecologists, an inescapable consequence would be the harsh judgement that unreplicated studies are woefully inadequate as tests of meaningful hypotheses. If this evaluation (which is consistent with many other areas of experimental biology and other sciences) were used, unreplicated studies should be ignored in any assessment of the value of field experiments on competition. To do so here would eliminate some 28 (29%) of the 95 papers examined, whether in whole or in part as having inadequate levels of replication to allow meaningful conclusions to be drawn from some or all of the experimental work. This is not a dissimilar

proportion of the total from that found by Hurlbert (1984) to contain pseudoreplication in some form or another (although the analyses are not entirely independent as some studies are common to both surveys).

11.3.2 Problems with controls

In addition to possible problems associated with experimental fences, cages, installations etc., some studies lacked sensible controls for parts of the manipulations. DeLong (1966), for example, examined interactions between house mice (*Mus*) and *Microtus californianus* on two grids. Both grids were then supplied with extra food, and *Microtus* were removed from one. In addition to the lack of replication, there were no controls for the experiment, and the results can only be used to infer processes where food supplies are artificially supplemented. Price (1978) made a detailed examination of habitat utilization by four species of rodents in a sampling area. Then, utilization of habitat was studied in experimental enclosures containing eight introduced individuals of a single species (Treatment 1 in her paper). She concluded that habitat use was less specialized than that in the intact community, because competitors were absent. As well as the intrinsic differences between the natural area sampled and the area used for the enclosures, and the possible artefacts of behaviour due to being in enclosures that might restrict the large scale movements of rodents, the experimental animals had also been trapped, handled and introduced into new, unfamiliar areas. Any of these three, uncontrolled variables, rather than the absence of other species, could explain the observed difference in behaviour compared with the natural community.

Lack of controls for the experimental disturbances of transplanted animals was found in several other studies. Pontin (1960, 1969) transplanted nests of ants (*Lasius niger* and *L. flavius*) to determine the nature of interactions with surrounding colonies. There were, however, no attempts to control for the effects of three types of experimental disturbances inevitable in such experiments. Firstly, there is the disturbance of the surrounding colonies due to digging out a space to insert a transplanted nest. Secondly, the transplanted ants are, presumably, seriously disturbed (and perhaps made more agitated?) by the physical disruption of the nest. Finally, the transplanted ants may behave differently as a result of being translocated to a new piece of habitat. Any of these disturbances, rather than just the presence of a putative competitor, could explain subsequent changes in surrounding nests. Controls for such disturbances are straightforward — comparisons could be made with areas in which nests were dug out and replaced (so that nests are disturbed but there is no translocation and no change of species); areas

where a nest is removed but replaced with a nest of the same species (the nest is disturbed and translocated, but there is no change of species) and the usual experimental plot where a nest of another species is introduced (disturbed, translocated and a change of putative competitor). Using this protocol, the various experimental disturbances could be separated from the effects of the primary experimental variable — the introduction of a different species.

In contrast to very active animals, such as ants and rodents, it is not likely that experimental disturbance would seriously alter the behaviour, rate of predation etc., of sluggish animals, such as starfish (Menge 1972). Nevertheless, the lack of controls for disturbance confounds any comparison of experimental and other types of control areas.

Other experiments lacked different types of controls. Fowler (1981) removed several species of plants from experimental plots in oldfields, and Gross and Werner (1982) dug up the soil in some areas before planting seeds. In neither case was there a control for the disturbance of the surface of the soil as opposed to the removal, or absence, of competing species. Seeds in recently disturbed ground might germinate better than those in undisturbed plots, regardless of the presence of other plants. If this were the case in either of the cited studies, removal, or absence, of potential competitors would, incorrectly, appear to have caused greater germination.

In some studies, attempts to provide proper controls for the experimental manipulations failed. Thus, Kroh and Stephenson (1980; experiment 2) analysed interactions among plants, by examining single plants (species A) surrounded by six individuals of the same (A) or one of three other species (B, C, D). Control plants were isolated with no surrounding individuals. Unfortunately, the lack of surrounding plants caused increased stress on the control plants, because they were no longer protected from wind and other physical factors.

It appears that controls for various necessary manipulations are often lacking, or inadequate. This compromises the value of any experiment by allowing several possible interpretations to be made about differences among experimental treatments. Confounding variables must be removed from consideration, by the establishment of appropriate sets of controls, even though these may be more unwieldy than the experimental manipulation itself (Connell 1974). Controls to remove confounding variables are essential, even if this causes changes in the circumstances under which the hypothesis of competitive interactions is tested. If disturbance of some experimental animals is necessary to complete an experiment (for example, to introduce members of a competing species into some experimental areas), then control groups of similarly disturbed animals that have not had the competitor introduced will also be necessary, to ensure that differences between treatments are

not due to handling and other types of disturbance, rather than the introduction of a competitor. If it is not possible to examine true controls (i.e. undisturbed animals) the nature of the experiment is such that the hypothesis being tested only applies to competitive interactions among disturbed (not necessarily *natural*) groups of organisms. This alteration of the hypothesis will, of course, reduce the generality of the results to those circumstances investigated. This, in turn, will limit the validity of extrapolations to entirely natural circumstances. More care in the establishment of proper controls for experimental procedures will make some experiments more complicated. It will, however, also increase the degree to which results of such experiments can be considered relevant to natural populations.

11.4 Confounded comparisons and other problems

11.4.1 Confounding in designs of experiments

Several different experimental designs have been used to investigate interspecific competition. Of these, two major types have been used extensively. The first is known as a replacement series, or reciprocal α design (de Wit 1960; de Benedictus 1974) and consists of a series of experimental plots, all at a fixed total density, but made up of different proportions of two species, ranging from a pure 'stand' of one through various mixtures to a pure 'stand' of the other. This design has been used commonly by plant ecologists. [See the extensive examples in Trenbath's (1974) review.] It has also been used in studies of animals [e.g. protozoa by Gill and Hairston (1972), freshwater bivalves by Mackie *et al* (1978)]. It is, however, often difficult to interpret the results of such experiments unambiguously, because any comparisons between experimental plots must involve two differences — the greater density of one species and the simultaneously reduced density of the other species. If species suffer from intraspecific in addition to interspecific competitive interactions (as they almost invariably will — see later), such experiments are of dubious value for interpreting processes of competition. In fact, some competition between two species might be completely masked by intraspecific effects of equal magnitude so that no change from treatment to treatment would be recorded when members of one species are replaced by members of another.

The alternative design, sometimes termed a 'mechanical diallel' (Putwain & Harper 1970) also makes comparisons between pure 'stands' of each species and mixtures of the two. Here, mixtures may be created by adding some number of another species to the individuals in a pure stand [a 'synthetic' design in Putwain and Harper (1970)] or pure stands

may be created by removing members of one species from a mixture (Putwain and Harper's 'analytic approach'). In either case, the total density of all individuals (i.e. of both species combined) will be different between single species and mixed plots (Table 11.1b). Comparisons between such plots, provided common sense has been used, will be unconfounded. Any difference between members of species A at some chosen density when alone, and when members of species B are present can, presumably, only be due to the presence of species B. Care must be taken, however, to ensure that the experimental design is not inadvertently confounded. In several studies, the general design of a 'mechanical diallel' was followed, but densities of each species when alone were different from those when the two species were together. Thus, Bertness (1981a) compared rock pools with 200 *Clibanarius* (hermit crabs), pools with 200 *Calcinus* and pools with a mixture of 100 of each species. Jaeger (1971) and Werner and Hall (1977) created precisely the same difficulties in experiments with salamanders and fish, respectively. Haven (1973) removed limpets (*Acmaea digitalis* and *A. scabra*) to make experimental plots with single species, but caused confounded comparisons between plots because of very different starting densities of the two species (Underwood 1976). In these studies, comparisons between areas with a single species and with a mixture were therefore confused because the addition of a putative competitor was confounded with the simultaneous decrease in density of the target species (Table 11.1a).

Table 11.1. Confounded and valid designs of experiments to investigate interspecific competition. (a) Confounded comparisons. There are two differences between individuals of species A in treatments 1 and 3 — the intraspecific change of density of species A and the absence/presence of species B. (b) Unconfounded comparisons. There is only one difference between individuals of species A in treatments 1 and 3 — the absence/presence of species B.

	Experimental treatment		
	1	2	3
(a)			
Area (units)	5	5	10
Number of organisms	10 A	10 B	10 A + 10 B
Density of species A per unit area	2	–	1
Density of species B per unit area	–	2	1
(b)			
Area (units)	5	5	5
Number of organisms	10 A	10 B	10 A + 10 B
Density of species A per unit area	2	–	2
Density of species B per unit area	–	2	2

A more useful, and more validly interpretable design is of the kind illustrated in Table 11.1b. The hypothesis proposed is that there is competition between species A and B, and thus the presence of B will deleteriously affect members of A, and vice versa. A sensible null hypothesis to test is that the addition of some number of B to some number of A will have no effect compared with control groups of that number of A when on their own. This null hypothesis cannot be tested by the confounded design in Table 11.1a. It is not obvious what null hypothesis could be tested by that design. The unconfounded designs (as in Table 11.1b) have been used successfully by many authors (e.g. Wilbur 1972; Lynch 1978; Hairston 1980b, 1981; Peterson & Andre 1980; Wise 1981a, 1981b; Creese & Underwood 1982).

Other forms of confounding were inevitable in some designs of experiments. Grant (1971), for example, discussed the inherent confounding of his experimental manipulations of densities of rodents with differences in habitat from one enclosure to another and one time to another. Dhondt and Eyckerman (1980) demonstrated different patterns of utilization of experimental nest-boxes by blue tits before and after the openings of the boxes were made smaller, thus preventing the larger great tits from entering them. It is difficult to attribute this finding unequivocally to the removal of competition from great tits; the blue tits rarely used the boxes when the entry holes were large before the experiment began. The experiment is confounded because of the possibility that reducing the entry hole made the boxes more suitable for blue tits regardless of the concomitant exclusion of the larger species. Miles (1972) compared the germination of seedlings in the presence and absence of existing coverage of plants. Seeds of each species were planted in separate, experimentally bared plots. On covered plots, however, seeds of four species were scattered together. Differences in germination of any species between bared and covered plots might be due to the presence of seedlings of the other species, rather than the hypothesized pre-emption of space by older plants.

11.4.2 Other problems with experiments

An extremely common problem with experimental manipulations of active, wide-ranging animals is that of maintaining the treatment after the start of the experiment. Schroder and Rosenzweig (1975) and Abramsky *et al* (1979) have described the very rapid re-invasions of rodents after they had been removed from experimental grids. Pajunen (1982) described rapid invasions of adult corixids, *Callicorixa*, into pools supposed to contain only *Arctocorisa* (although there were no data of any kind in that paper). Sometimes, the competitors can be manipulated, but the resource

is mobile. To determine any effect on the behaviour and distribution of ducks, Eriksson (1979) divided a lake into two with a net, and introduced fish into one half. The food of fish and ducks (adult and larval insects) could, however, move between the two halves, and may not have been reduced in abundance where fish were present.

Great spatial and temporal variations in numbers of a population also cause problems for small-scale or short-term experiments. Thus, in Tinkle's (1982) experiments on lizards, there were large changes in numbers of *Sceloporus undulatus* in the control area where two other species were present during the study. Such variations make it very difficult to detect any differences between areas with and without competitors, unless enormous numbers of replicate areas, or very long time courses of experiments are available. Such variations can, however, be expected for many organisms and should lead to variations in the intensity or importance of competitive interactions from time to time and place to place (Lynch 1978; Hils & Vankat 1982; Underwood *et al* 1983).

From this brief account of some of the problems with experimental procedures, it is clear that many studies of competitive interactions are going to lead to, at best, ambiguous results. Greater attention to the design of the experiments will remove obvious sources of confounding. Other problems are, however, unavoidable, although they might be detected by short term pilot studies in advance of committing large amounts of time and energy. As the whole purpose of experimental manipulations of hypothesized competitors is to produce unequivocal conclusions about the presence, intensity or nature of competition, many experiments must be considered far more preliminary than has been the case to date. It is doubtful that some of the confounded comparisons in the literature are of real value as demonstrations of competition as an important process. More care in the subsequent interpretation of such experiments might lead to better designs for future studies, rather than simple uncritical acceptance of the results regardless of their inherent difficulties.

11.5 Inter- and intraspecific competition?

11.5.1 Why study intraspecific competition?

There are four major reasons why interpretations of studies of interspecific competition would benefit from simultaneous examination of intraspecific interactions among the individuals of each of the competing species. First, the detection of intraspecific competition is a good starting basis from which to build hypotheses about interspecific competition, because the

similarities of requirements for essential resources must be greater (or, at the very least, as great) among members of the same species than among members of different species. Thus, when resources are in short supply, intraspecific interactions must be occurring, if interspecific competition is also to occur. The point was raised by Darwin in *The Origin of Species*: "struggle for life most severe between individuals and varieties of the same species" (heading on p. 59 of the 1882 edition). This has been discussed by many authors (e.g. Pontin 1969) and was stated explicitly as criterion 3 in Reynoldson and Bellamy's (1970) discussion of necessary and sufficient conditions for demonstrating the existence of interspecific competition in natural populations. Yet it is often forgotten, and many experiments on interspecific competition are done without any investigation of the effects of members of the same species. Among other problems caused by this omission were the confounded comparisons in experiments described previously.

Secondly, intraspecific competition is clearly important in the development of theory about competitive interactions and subsequent development of general community theory. [For examples of Lotka–Volterra competition models, see Williamson (1972), Krebs (1978), Lawton & Hassell (1981) and Schoener (1983b).] Thirdly, the magnitude and intensity of interspecific competition detected in experiments will probably depend greatly on the starting densities of the species investigated, relative to their carrying capacities, because of intraspecific as well as interspecific interactions. [See the example of a nonlinear analysis in Schoener (1983), and the discussion in Connell (1983).] Fourthly, the outcome of interspecific competitive interactions, in terms of distribution and abundances of competitors, or in terms of structure of guilds or communities, will often depend on the relative importance of intraspecific in addition to interspecific interactions. Thus, an inferior competitor amongst intertidal limpets was able to coexist where intraspecific competition amongst members of the superior competitior was sufficiently intense to keep its density below the levels at which it could exclude the inferior species (Creese & Underwood 1982). Pontin (1969) discussed the regulatory role of intraspecific competition in returning competing species back to a stable equilibrium rather than a situation where the inferior species was excluded.

Sensible designs of experiments for the simultaneous analysis of inter- and intraspecific competition have been used by several authors (Wilbur 1972; Underwood 1978; Waser 1978; Seifert & Seifert 1979; Kroh & Stephenson 1980; Fonteyn & Mahall 1981; Brown 1982; Creese & Underwood 1982). Such experimental protocols are much more informative and reliable than indirect estimations of intraspecific interactions (e.g. Abrams 1981a).

11.5.2 *Designs of experiments on inter- and intraspecific competition*

The general design of experiments to investigate inter- and intraspecific competition between two species is illustrated in Table 11.2. Extensions to more than two species may be found in a number of studies on intertidal gastropods by Underwood (1978), Creese and Underwood (1982); salamanders by Wilbur (1972); cladocerans by Smith and Cooper (1982) and insects in pitcher plants by Seifert and Seifert (1979). Powerful statistical analyses of such experiments are available, particularly where orthogonal contrasts can be made simultaneously amongst several sets of treatments (e.g. Wilbur 1972; Underwood 1978; Creese & Underwood 1982). Experiments are described where intraspecific components of competition are investigated at only two densities of each species (Table 11.2a), and where several experimental densities are used (Table 11.2b). The advantage of the latter type of experiment is that the relationship between competitive effects and starting densities can be unravelled. [See the nonlinear plots in Schoener (1983b).] There are, however, problems where the natural, or otherwise relevant, densities of two species are not similar (see below).

Table 11.2. Designs of experiments to investigate intra- and interspecific competition between two species. (a) Two experimental intraspecific densities: both species at same density. (b) Three experimental intraspecific densities: both species at same density.

(a)

Density	Experimental treatment				
	1	2	3	4	5
Species A	10	20 (= 10A + 10A)	10	–	–
Species B	–	–	10	10	20 (= 10 B + 10B)
Comparison:		Intraspecific competition 2 vs 1 (A on A) 5 vs 4 (B on B)		Interspecific competition A's in 3 vs 1 (B on A) B's in 3 vs 4 (A on B)	

(b)

Density	Experimental treatment								
	1	2	3	4	5	6	7	8	9
Species A	10	20	30	10	10	20	–	–	–
Species B	–	–	–	10	20	10	10	20	30
Comparison:	Intraspecific competition 3 vs 2 vs 1 (A on A) 9 vs 8 vs 7 (B on B)			Interspecific competition A's in 5 vs 4 vs 1 (B on A) B's in 6 vs 4 vs 7 (A on B)					

Interpretations of interspecific interactions would benefit from simultaneous examination of intraspecific interactions as pointed out by Connell (1983). Relatively few studies of the total examined included any attempt to investigate effects of intraspecific increases in density on each species. This is again surprising given the long time-scale over which authors have urged consideration of intraspecific effects as a necessary prerequisite or corequisite to analyses of interspecific competition (Reynoldson & Bellamy 1970). The situation has changed very little since Pontin's (1969) statement that "this point [i.e. the importance of intraspecific interactions] continues to be overlooked in discussions of interspecific competition".

11.6 Symmetry and asymmetry in interspecific competition

11.6.1 Three types of symmetry

It has recently been fashionable to discover that competition between two species is rarely symmetrical — usually one species has greater effects on a second species than it experiences from the second species (e.g. Lawton & Hassell 1981; Connell 1983; Schoener 1983b). In terms of Lotka–Volterra theory, such asymmetry occurs whenever the competition coefficient (α_{ij}'s) are unequal. In the examples of experimental designs in Table 11.2, symmetry exists when the effect of species A on species B equals the effect of species B on species A. Lawton and Hassell (1981) concluded that asymmetrical competition was the norm for insects. Schoener (1983b) emphasized the importance of asymmetrical competition and suggested several possible processes to account for superiority by one species over another. I will not discuss here the importance of asymmetrical competition to the development of theory. Instead, I suggest that symmetry/asymmetry exists in three different forms, and each has intrinsic worth for understanding the role of competition among species in natural communities. First is the interspecific symmetry considered by previous workers (i.e. the degree of similarity of effect of each species interspecifically on the other species).

Of similar importance, however, is the degree of symmetry of interspecific versus intraspecific effects of individuals of a species. In the designs in Table 11.2, this symmetry exists only if the intraspecific effect of species A on species A is equal to the interspecific effect of species A on species B. This is important for predicting the effects of changes of numbers, fecundities, weights, survival, etc. of any species when new individuals of either species invade, arrive, recruit, or are experimentally introduced. Thus, if species A is asymmetrical by having a greater intra-

specific competitive effect than its interspecific effect on B, an experimental increase in densities of species A into an area where A and B coexist will cause greater negative changes in members of species A than of species B (and could therefore actually have positive effects on species B if the resulting deleterious effects on A were of sufficient magnitude to reduce the overall resulting density of A in the area). For example, consider the resource of space on a rocky shore, and the situation when space in one area is occupied by members of species A and another area is occupied by members of species B. To make reliable predictions about future events in either area, when recruitments of further A individuals occur, will depend entirely on the relative competitive abilities of A on the occupant individuals of the species in each area. If A has a greater effect on existing A than on individuals of B, then recruitment of further A into each area will cause greater change where A occupies the space than where B occupies the space.

The third type of symmetry/asymmetry also involves comparisons of inter- and intraspecific competitive interactions. This is the comparative effect of members of species A on other A versus the effects of species B on species A. Knowledge of this form of symmetry would enhance predictive power about invasions or recruitments of species. Again, consider the resource of space on a rocky shore, and the situation when space in one area is occupied by members of species A. Future events in this area, if recruitments of further A or members of some other species B might occur, will depend entirely on the relative competitive abilities of the two species on the occupant members of species A. Thus, asymmetrical effects (A on A not equal to B on A) will cause different outcomes depending upon whether individuals of species A or B arrive first.

These three types of symmetry and asymmetry have not yet been considered for the complete analysis of any interaction, mostly because intraspecific effects of competing species are often ignored in many studies. They will not be considered further here, but, if one form of symmetry is important to the development of a complete understanding of competitive interactions, so are the other forms. Differences in the relative strengths of intraspecific interactions within two species might go a long way towards explaining asymmetrical interspecific competitive abilities, but have not been considered in this context in recent discussions (Lawton & Hassell 1981; Connell 1983; Schoener 1983b). One example is the inferior competitive ability of the limpet *Cellana* compared with the snail *Nerita* on rocky shores in New South Wales (Underwood 1978). *Cellana* also suffers much more intense intraspecific competition than does *Nerita* (Underwood 1976, 1978; Creese & Underwood 1982).

11.6.2 Experimental designs for the analysis of symmetry

The types of experiments outlined earlier (Table 11.2) would produce all the information necessary for complete analyses of all forms of symmetry. There is, however, a major problem when the densities of the competing species are not the same. Consider a simple example of competition between two species that are exactly equal (symmetrical) in their competitive effects (both inter- and intraspecifically). For the purposes of illustration, assume that members of both species consume all resources in proportion to their densities (thus, all food resource in an area is divided equally amongst all individuals, of whatever species, present in the area). Finally, consider that the weight of tissue of each individual increases during a given period of time by half the excess amount of food consumed over that required for maintenance (this is not realistic, but, to keep the example simple, no adjustments are made for changing weights of the animals). Similarly, if insufficient food is available for maintenance, tissue weight declines by half the deficiency of maintenance requirement minus amount consumed. Also, assume that densities of the two species are at some different number per unit area (but combined density represents the carrying capacity) when an experiment is planned. This situation is illustrated in Table 11.3, under experiment 1, where the experimental density is eight individuals of species A and five of species B. Each then gains exactly the maintenance requirements of food, and neither grows nor decreases in weight. Relative to the experimental plots where the other species has been removed, however, the animals in this plot will not weigh as much at the end of the experiment. The differences in weights between members of each species when alone and when together will be used by the experimenter as an estimate of the competitive effect of each species on the other. Apparent interspecific asymmetry will be inevitable in such an experiment. Note that the relative effect per individual of species A is to decrease the weight of individual B by 0.05 units (i.e. 0.40 divided by the 8 A in the mixed treatment). The effect of individual B is to reduce the weight of individual A by 0.03 units (0.16 divided by 5). This asymmetry is entirely artificial and occurs simply as the result of using different densities of the two species in the mixed treatment.

In experiments 2 and 3, each species is used at the same density (5 and 8 per enclosure in the two experiments; Table 11.3). Note that different relative effects of the two species are estimated in the two experiments: B has more effect on A when there were 5 A than when there were 8 A per enclosure. This is consistent with the non-linear analysis described by Schoener (1983b). In each of these experiments, however, the interspecific effects of each species on the other are symmetrical, as they should be from the situation modelled. Interspecific

Table 11.3. Symmetry and asymmetry in interspecific competition as a function of experimental design. Individuals of species A and B are equivalent (symmetrical) competitors needing 0.5 units of resource for maintenance. Each gains or loses weight as half the excess or shortage of resource compared with that needed for maintenance. The natural mean density in the area is 8A and 5B per unit area, and there are 6.5 units of resource available.

	Experimental density		
Experiment 1	8A	8A + 5B	5B
Consumption by individual A	0.81	0.50	–
Change of weight of A	+0.16	0	–
Change relative to 8A alone	–	−0.16	–
Consumption by individual B	–	0.50	1.30
Change of weight of B	–	0	+0.40
Change relative to 5B alone	–	−0.40	–
Experiment 2	5A	5A + 5B	5B
Consumption by individual A	1.30	0.65	–
Change of weight of A	+0.40	+0.08	–
Change relative to 5A alone	–	−0.32	–
Consumption by individual B	–	0.65	1.30
Change of weight of B	–	+0.08	+0.40
Change relative to 5B alone	–	−0.32	–
Experiment 3	8A	8A + 8B	8B
Consumption by individual A	0.81	0.41	–
Change of weight of A	+0.16	−0.05	–
Change relative to 8A alone	–	−0.21	–
Consumption by individual B	–	0.41	0.81
Change of weight of B	–	−0.05	+0.16
Change relative to 8B alone	–	−0.21	–

symmetry can only be properly detected when the same densities of the competitors are examined — because of the simultaneous effects of intraspecific competitive interactions among the individuals of each species.

It is not clear how much such confusion might have influenced conclusions about symmetry in previous analyses (Lawton & Hassell 1981; Connell 1983; Schoener 1983b), but all of these reviewers found a preponderance of asymmetry, but did not discuss the relative number of each species in the experiments. Some of the experiments (e.g. Peterson & Andre 1980; Williams 1981) certainly used different numbers of the competing species. Others did not (e.g. Brown 1982; Smith & Cooper

1982). Conclusions about the prevalence of symmetry and asymmetry in the literature are premature until the potential effects of experimental designs have been considered carefully. This does not seem to have been the case to date. As suggested earlier, the analysis of intraspecific competition in relation to interspecific competition is also made more difficult when different densities of two species are to be used to cause experimental mixtures.

In conclusion, symmetry, in three different forms, is important to allow realistic evaluation of the nature and importance of competitive interactions. Experiments must be more carefully designed with hypotheses about symmetry firmly in mind before any useful conclusions could be drawn about the existence of, or explanations for, symmetrical or non-symmetrical competitive abilities. Hypotheses about symmetry must, presumably, relate to the competitive effect per individual of competing species. Therefore, before any conclusion can be drawn from the literature, some manipulation of experimental results is necessary to correct for the numbers of each species, where densities of the competing species were different.

11.7 Conclusions

The major conclusion that emerges about experimental analyses of interspecific competition is a bewilderment that there are so many problems of design, implementation and interpretation in so many experiments. This does not include problems with such things as the statistical analyses (Underwood 1981; Hurlbert 1984) that have yet to be discovered. All in all, the message for future historians of ecology is 'do not look — spare yourselves'. After all, the bones of the principles of experimental design are at least 50 years old, but have not yet been fleshed out in the corpus of ecological investigation. This is despite the very recent development of ecological experiments on interspecific competition (Schoener 1983b). There are far too many problems with replication and controls, with unclear relationships between hypotheses and tests, and with confounded comparisons for us to be complacent. With Hurlbert (1984), I bemoan the poor state of experimentation, and suggest most forcefully that Schoener (1983b, p. 241) *was* premature to try his analysis, and that Connell (1983) was being excessively optimistic to assume that it was only when field experiments were *few* that the "tendency was to use this small amount of information uncritically".

One very clear conclusion is apparent — when evidence for competition is sought, seemingly any evidence will do. Before the trend to do experiments, competitive interactions were often invoked to explain many patterns in nature, without consideration of other equally tenable hypo-

theses. Now, regardless of the rigour, or lack of it, of an experimental design, competitive interactions are inferred from most experiments — even when other equally tenable hypotheses have not been eliminated by the experiment. A particularly obvious example was provided by De Long (1966, p. 481) who found a negative correlation between the densities of house mice and meadow mice in his sampling areas. He then manipulated the densities of *Microtus* "to see if the negative correlation between the densities of the two species was due to interference by *Microtus*, or was merely spurious". It could, of course, have been real — but due to other causes than competition. Other workers have not been so obvious about ignoring other possible models and hypotheses. If this state of affairs continues, there will have been little point in the experiments. Many will have been planned, carried out and interpreted solely as confirmations of existing hypotheses. [Aristotelian experimentation as defined in Medawar (1969), see also the thoughtful discussion of Dayton (1979).]

The present analysis confirms the worst fears of Hurlbert (1984). To paraphrase him, poorly designed and ill-conceived experiments on competition have literally flooded the literature during the last few years. This parlous situation can, of course, only come about with the active connivance of grant agencies, referees of papers, editors of journals, chairmen of scientific meetings, referees for promotion, selection committees for employment, compilers of reviews of the literature and the general ecological community. Why are we so afraid of constructive criticism about experimental work? Have we nothing new to learn about nature, and therefore know all there is to know about experiments? Perhaps it is time for experimental ecology to adopt a harsher judgement of poor work and reject it as inadequate, sloppy and open to numerous alternative explanations and conclusions than those reached about competition. Many of the problems could have been avoided by better designs of experiments, and therefore by more critical appraisal during the planning stage (by various of the categories of person listed previously). As a step to rid ecological theory of the encumbrance of dealing with inadequate experimental tests of former pieces of theory, and in the hope of moving towards an improved future set of experiments (rather than continued uncritical acceptance of the present examples followed by inadequacy becoming the norm for future work), the time is at hand to reject all unreplicated, uncontrolled, confounded, or otherwise inadequate experiments from further consideration. As a preliminary move towards this, papers that have offered sensible, competent or valuable description, discussion or interpretations about the design, planning and analysis of experiments on competition in the field have been listed (Table 11.4). Most of these papers probably also contain faults, but each has something to offer. This is obviously a personal choice amongst the papers sampled — many other papers are also valuable, but were not included in this

Table 11.4. Studies that provide useful experimental designs, discussion or interpretations of field experiments on competition.

Study organisms	Author
Desert plants	Friedman (1971); Friedman & Orshan (1974), Friedman *et al* (1977), Inouye (1980)
Old-field plants	Pinder (1975), Allen & Forman (1976), Gross (1980), Hils & Vankat (1982)
Heath plants	Miles (1974)
Prairie plants	Petranka & McPherson (1979)
Grassland plants	Putwain & Harper (1970)
Winter annuals	Raynal & Bazzaz (1975)
Jack-pine	Brown (1967)
Paramecium	Gill & Hairston (1972)
Pond-snails	Brown (1982)
Sea-pansy	Kastendiek (1982)
Fouling organisms	Sutherland (1978)
Anemones/barnacles, etc.	Taylor & Littler (1982)
Bivalves	Peterson & Andre (1980)
Marine gastropods	Creese & Underwood (1982), Underwood (1978)
Urchins/fish	Williams (1981)
Bees/flies	Morse (1981)
Beetles	Wise (1981a)
Spiders	Wise (1981b)
Salamanders	Hairston (1980b, 1981), Wilbur (1972)
Lizards	Smith (1981)
Rodents/fungus	Inouye (1981)
Rodents	Joule & Jameson (1972), Redfield *et al* (1977)

survey. I suggest these as the examples to be consulted early in the planning of future experimental work. I also urge all ecologists to read more about the design of experiments, and to consult cautionary reviews such as that by Hurlbert (1984).

The resources for which competition might occur and the mechanisms used by potential competitors to gain adequate amounts of resources are well reviewed by Connell (1983) and Schoener (1983b). These reviews and the present discussion confirm that competitive interactions can be investigated, interpreted and understood by use of manipulative field experiments. Many of the examples in Table 11.4 demonstrate this (although some demonstrated, equally validly, that competitive interactions were not important in some systems). The validity of the experimental procedures, however, depends entirely on the care with which the hypotheses being tested are defined. How closely the experimental conditions represent reality depends entirely on a clear understanding of the

reality under investigation. The adequacy of controls for experimental manipulations is entirely determined by the experimenter. The need for replication in any field experiment is paramount — numerous sources of natural variation are already well-documented for most types of habitat and organisms. When questions are addressed about how widespread competition is in some ecosystems, more replication in space and time is needed. Otherwise, the question cannot be answered, even by the most elegant experiment at one site or time. The criticisms offered here are not to dissuade ecologists from attempting experiments, but to provide a framework in which improved methods for interpreting nature can be developed.

The papers listed in Table 11.4 discuss experimental designs, adequacy of controls and replication with respect to a wide variety of organisms, habitats and interesting hypotheses about competitive interactions and their role in structuring natural communities. There exists no more valuable guide to the methods available to overcome problems with field experiments.

Competition is undoubtedly important as an organizing force in the structure of communities and in the evolution of the organisms in many communities, as revealed by the syntheses of Connell (1983) and Schoener (1983b). Good, unequivocal evidence of this has been provided by many field experiments. The less valuable experiments are not needed to demonstrate the generality of competitive interactions in many types of environment.

Nevertheless, the prevalence of interspecific competition is not as well understood as suggested by the large number of experimental papers currently available in the literature. This situation will only be improved by a series of better experiments, not by further analysis of the existing ones. The smaller number of studies that have produced unequivocal results demonstrates that it is possible to do good field experiments on competition, and that such experiments still provide a powerful tool for unravelling complex biological communities. The instruments are, however, badly blunted by consistent, widespread, incorrect usage.